Numerical Weather Prediction

Verification Report for Rainfall Downscaling using the ECMWF Ensemble Prediction System

Forecasting Research Technical Report No. 445

Sarah E John

email: nwp_publications@metoffice.com

©Crown Copyright

Verification Report for Rainfall Downscaling using the ECMWF Ensemble Prediction System

Sarah E John. First version 07/07/04. Revised version 02/11/04

Abstract

Site-specific quantitative precipitation forecasting using numerical weather prediction (NWP) models is problematic. Sub-grid-scale processes and local site biases determine that the grid box average precipitation values predicted by NWP models can be considerably different from the observations recorded at individual sites. A method of statistically downscaling NWP model precipitation fields to generate reliable probability of precipitation forecasts is clearly desirable. To achieve this aim precipitation should be modelled in an appropriate manner. It will be argued that the correct method to model the probability distribution of precipitation is as a mixture of the wet and dry event probability density functions.

A method of statistically downscaling model precipitation forecasts to individual sites in the UK has been developed, using the control member of the ECMWF Ensemble Prediction System (EPS). The downscaling method has been designed to estimate the sub-grid-scale uncertainty affecting the precipitation forecast for an individual site but takes no account of the synoptic-scale evolution of the atmosphere. Assuming that the EPS represents the synoptic-scale uncertainty the statistical downscaling is applied to each member of the EPS to provide a more complete description of the uncertainty. The downscaling system was tested for nine months from 1st January 2003 to 30th September 2003. Verification results indicate that the downscaled EPS probability forecasts exhibited maximum improvement over raw EPS probability forecasts at the 0.1mm (wet/dry) threshold.

1. Introduction

Precipitation is one of the most difficult parameters to forecast accurately. Results from the World Meteorological Organisation (WMO) Working Group on Numerical Experimentation indicate that despite frequent upgrades, the leading global numerical weather prediction (NWP) models still experience difficulty in making accurate quantitative forecasts of precipitation (Ebert et al., 2003). Site-specific, quantitative precipitation forecasting from NWP models is even more problematic. Model output can be directly interpolated to make predictions at specific sites; however, model values are inherently grid-scale quantities and it is clear that model precipitation fields are much smoother than the spatially and temporally intermittent observed precipitation fields (Ebert et al., 2003). Convective parameterisation schemes attempt only to represent the mean behaviour of an ensemble of convective clouds and showers within a grid-box. Similar averaging of sub-grid-scale detail occurs with large-scale precipitation although to a lesser extent. Furthermore, the true resolution of a model is typically 4-5 grid lengths. Local site biases due to topographic variations (for example, proximity to large

bodies of water, large urban areas or orographic variations too small to be handled by the smoothed model orography) are not accounted for by NWP models. Hence, grid box average precipitation forecasts from NWP models can be considerably different from the corresponding observations at any individual site within the grid box.

Probabilities provide a concise and unambiguous method of expressing the uncertainty associated with forecasts (Murphy, 1978). The main source of probabilistic forecast information issued by the Met Office is the Ensemble Prediction System (EPS) run at the European Centre for Medium Range Weather Forecasting (ECMWF), fully described by Molteni et al., (1996) and Buizza et al., (1999 and 2000). A comprehensive range of probabilistic forecast products are available from the Predictability Visualisation (PREVIN) system run at the Met Office including a number of site-specific probability products, including probability of precipitation (Legg et al., 2002). The EPS is primarily designed to estimate the uncertainty in the large-scale evolution of the atmosphere and not the uncertainty due to unresolved scales. Therefore, a method of statistically downscaling EPS precipitation fields to generate reliable, site-specific probability of precipitation forecasts is desirable.

A statistical model has been developed to perform the statistical downscaling and this is presented in section 2. Section 3 describes the data used to estimate the statistical model and generate probability forecasts. Section 4 presents a comparison of the probability forecasts generated from simple climatological data, the raw EPS and the statistical downscaling method applied to both the control member of the ECMWF EPS and the whole EPS. Finally, section 5 contains some conclusions and recommendations.

2. The Statistical Model

A statistical model has been developed to perform the downscaling which is based on the argument that precipitation is a mixture of wet and dry events. The model states that bivariate precipitation data, consisting of forecasts, X, and corresponding observations, Y, can be modelled as a four-component mixture model of wet and dry event probability density functions. (Stephenson, 2003, personal communication)[1]:

$$f(x,y) = \pi_{dd}\delta(x)\delta(y) + \pi_{dw}\delta(x) f_y(y) + \pi_{wd}\delta(y) f_x(x) + \pi_{ww} h(x,y) \quad (1)$$

where

$\pi_{ww} = \Pr\{ X > \varepsilon \text{ and } Y > \varepsilon \}.$

$\pi_{wd} = Pr\{ X > \varepsilon \text{ and } Y \le \varepsilon \}.$

$\pi_{dw} = Pr\{ X \le \varepsilon \text{ and } Y > \varepsilon \}$

[1] Equation 1 provides a full description of the problem and not all the distributions contained in the equation are required to make a probability forecast; however, the extra information is included here for completeness

$$\pi_{dd} = Pr\{ X \leq \varepsilon \text{ and } Y \leq \varepsilon \},$$

$$\pi_{dd} + \pi_{dw} + \pi_{wd} + \pi_{ww} = 1$$

and

$$X \mid (Y \leq \varepsilon \text{ and } X > \varepsilon) \sim f_x(x)$$

$$Y \mid (X \leq \varepsilon \text{ and } Y > \varepsilon) \sim f_y(y)$$

$$X,Y \mid (Y > \varepsilon \text{ and } X > \varepsilon) \sim h(x,y)$$

ε is the threshold delineating a dry from a wet event. π_{dd} is the probability that both the forecast and the corresponding observations are dry, a dry-dry event; π_{dw} is the probability that the forecast is dry and the observation is wet, a dry-wet event; π_{wd} is the probability that the forecast is wet and the observation is dry, a wet-dry event; π_{ww} is the probability that both the forecasts and the observations are wet, a wet-wet event. $\delta(x)$ and $\delta(y)$ are delta functions representing the distributions of X and Y for dry-dry events. $f_x(x)$ is the marginal distribution of the forecasts for wet-dry events and $f_y(y)$ is the marginal distribution of the observations for dry-wet events. Finally, $h(x,y)$ is the joint distribution of precipitation for wet-wet events. Equation 1 is illustrated schematically in Figure 1. More details of the statistical model are provided in Appendix A and John (2003).

The statistical model is estimated using past forecasts and observations and is designed to estimate the uncertainty associated with unresolved scales. This method possesses similarities to the MOS technique developed by Glahn and Lowry (1972), where observations of precipitation were regressed onto forecasts of a number of atmospheric variables and systematic model biases were accounted for. The MOS system of Glahn and Lowry (1972) needs to be re-calculated at each forecast lead time to take account of the changing error characteristics of the model with lead time. One difference between the application of MOS and the statistical model described here is that the statistical model parameters are only estimated using short range forecasts of precipitation from the control member of the EPS. The statistical model parameters are not re-estimated at all forecast lead times otherwise errors in the forecast associated with the synoptic scale evolution of the atmosphere would become incorporated in the statistics. It is assumed that the synoptic scale errors present in short range forecasts (for example, T+12) from the control member of the ensemble are small compared to those at longer forecast lead times (for example, T+180) and that any systematic model biases captured by the statistical model are due to the unresolved scales only rather than model error at the synoptic scale. Therefore, we proceed on the basis that the different scales of uncertainty in a precipitation forecast, the synoptic scales and the unresolved scales, are represented by the ECMWF EPS

and the statistical model respectively. However, there will be a degree of overlap in practice.

The statistical model can be applied to a single deterministic model forecast (at any forecast range) to generate a forecast of the probability of precipitation exceeding a user-defined threshold. However, the statistical model is not explicitly designed to take account of the uncertainty associated with the synoptic evolution of the atmosphere; a single model precipitation forecast is just one possible solution to the prediction problem and will possess a degree of uncertainty associated with the synoptic scale. Given that each member of the ensemble is *a priori* equally likely, then the statistical model (estimated for the control member of the EPS) can be applied to each member of the EPS. An overall exceedence probability for the ensemble is then derived by averaging the probabilities obtained for each ensemble member. It is hoped that this will offer a more complete description of the uncertainty in a precipitation forecast including synoptic-scale and site-specific uncertainty. Therefore, probability forecasts are generated from the statistical model applied both to the control member of the EPS and to the whole EPS.

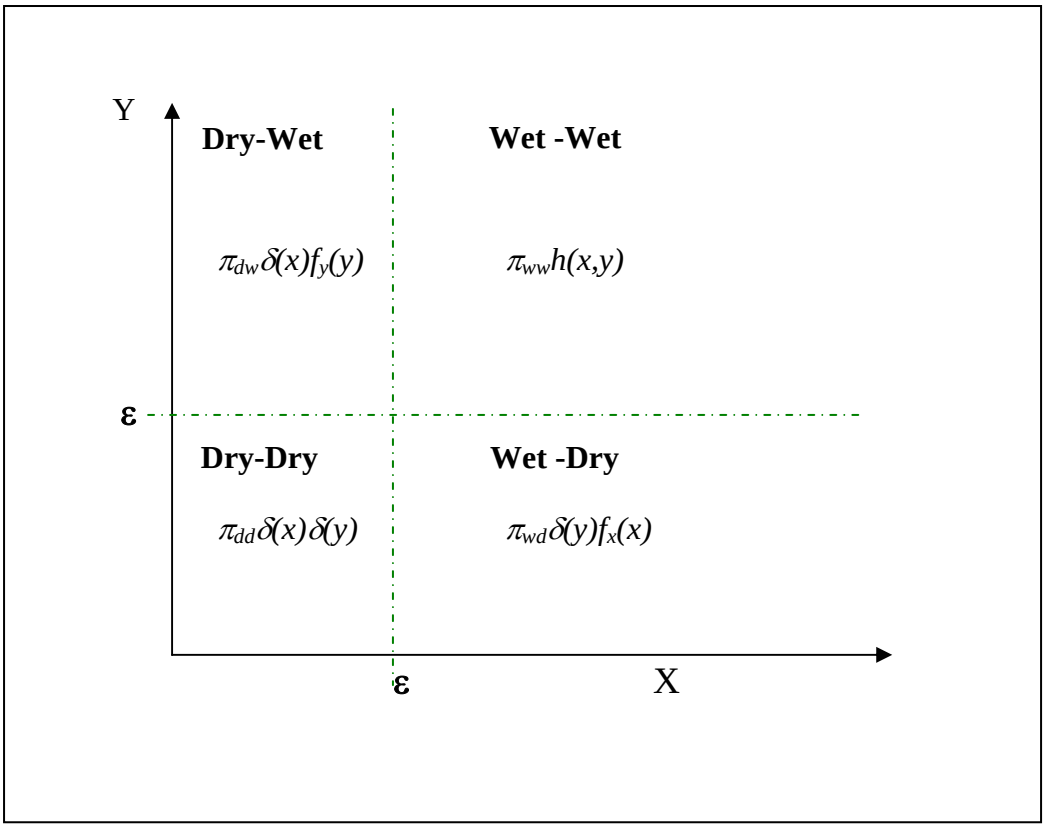


Figure 1. Schematic representation of Equation 1. Where Y are the observations, X the forecasts and ε the threshold delineating wet and dry events, all other symbols possess the same meaning as in Equation 1.

2. Data Used

The training data, used to estimate the statistical model parameters, consists of the model forecasts and corresponding gauge observations of the total precipitation accumulated in the twelve hours preceding 12UTC and 00UTC. The precipitation forecasts were obtained from the control member of the ECMWF EPS and bi-linearly interpolated from the four nearest model grid points to the chosen sites. At the time of this study the ECMWF EPS was run once a day at 12UTC; in order to calculate the statistical model parameters at 00UTC and 12UTC, T+12 and T+24 hour forecasts were used respectively. The statistical model has been estimated for five sites across the UK; Heathrow – 03772, London Weather Centre – 03779, Cardiff Weather Centre – 03717, Manchester Ringway – 03334 and Aviemore – 03063. Precipitation data are highly skewed. Therefore, prior to calculating the model parameters, the training data are transformed to become more normally distributed, using the Box-Cox transformation (Appendix A).

The statistical model is dependent on, and sensitive to, changes in the ECMWF forecast model and its error characteristics, as computing power increases and improved parameterisations are developed. Buizza et al., (2001) state that the ECMWF EPS underwent a major upgrade in November 2000, during which the horizontal resolution was increased from TL159 to TL255 (approximately 120km down to 80km grid-point spacing). Hence, the length of the developmental data sample was restricted to two years of data extending from 1st January 2001 to 31st December 2002 inclusive and no seasonal stratification was applied to the data. During the training period changes were made to the ECMWF system; mainly these have been within the 4D-VAR data assimilation code and a few changes to the model physics, including convection (Lalaurette, 2002a and 2002b). It should therefore be recognised that there may have been changes in the error characteristics of the precipitation forecast even within the developmental data sample. Data from 1st January 2003 to 30th September 2003 were reserved for verification purposes. Ideally the forecast system should be verified over a whole year's worth of forecasts, to avoid one season particularly biasing the results. Unfortunately lack of data determines that the independent data sample does not consist of a whole year's worth of data. Again, Andersson et al. (2003) state that the ECMWF model was upgraded on 14 January 2003 to include revised 4D-VAR code, new satellite data and changes to the cloud and convection schemes.

3. Verification Results

Brier (1944) recognised that when considering the problem of forecast verification, there is a need to distinguish between the scientific and practical objectives of a forecast. Verification measures that are suitable for scientific purposes are not necessarily appropriate to assess the economic value of the forecast. Forecasts have different attributes: reliability, resolution, discrimination and sharpness. Scientific verification is concerned with

understanding and improving the forecast process and requires more than just one or two summary scores. Therefore, the Brier score and Brier skill score (BSS), the reliability, the sharpness, the relative operating characteristic and the economic value of the forecast are considered here.

a) The BSS

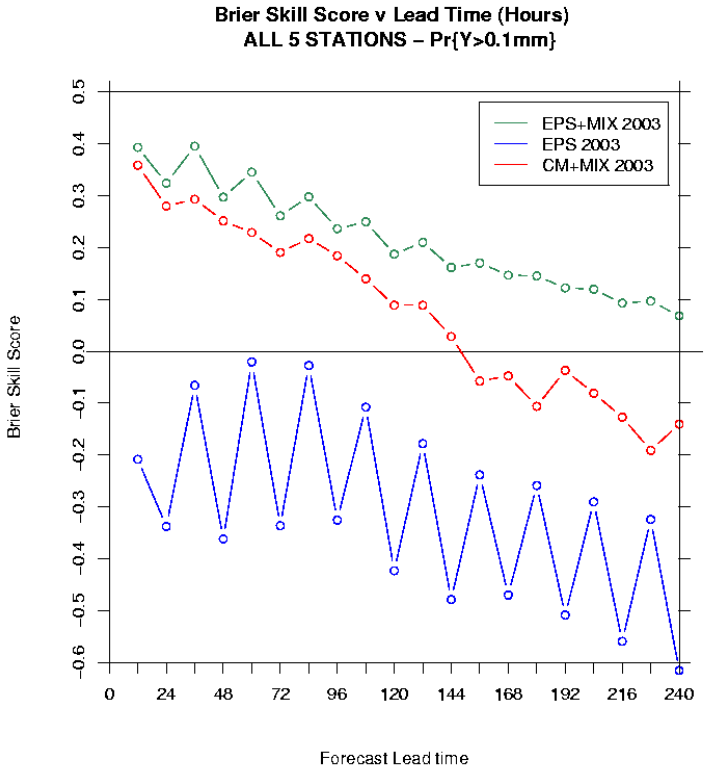
The Brier score (Brier, 1950) provides an overall measure of the accuracy of probabilistic forecasts of binary random variables. The performance of a forecast system relative to a base-line forecast system or climatology can be assessed using the BSS.

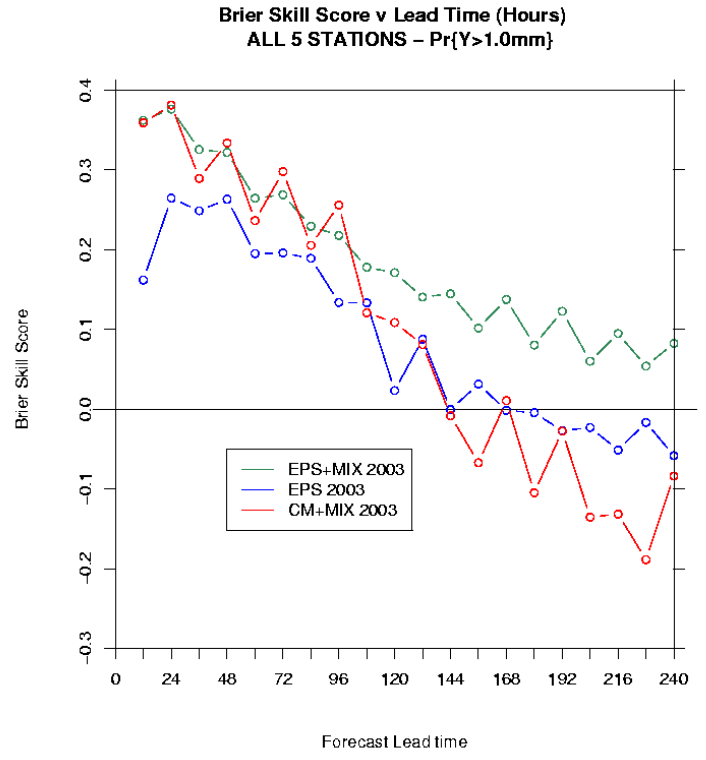
$$BSS = \frac{BS - BS_{ref}}{BS_{perfect} - BS_{ref}} = 1 - \frac{BS}{BS_{ref}} \qquad (2)$$

Where BS is the Brier Score, the subscripts *ref* and *perfect* refer to the Brier Score for the reference forecast system and a perfect forecast respectively. The BSS is positively orientated, ranging from minus infinity to unity. Maximum skill is obtained when the BSS is equal to one, a BSS of zero indicates the forecast skill is the same as the reference system; there is no lower bound to the BSS.

The reference forecast system was chosen to be a simple climatological forecast that issues the same fixed estimate of the probability of a given event each day. Although the climatological forecast is perfectly reliable (the probability of the event is equal to its frequency of occurrence) it takes no account of the current synoptic evolution of the atmosphere. Hence, if a forecast system is useful it should exhibit more skill than climatology. Climatological information for 12-hour accumulations of precipitation is not calculated routinely for the stations used in this study. Therefore, climatological forecasts (at 00UTC and 12UTC) for each station were estimated using the two years of observations held in the training data set.

Figure 2 presents results of the BSS versus forecast lead time, for all stations, for the thresholds 0.1mm, 1.0mm and 5.0mm. Three different probability forecasts are compared in Figure 2: probability forecasts from the raw EPS (blue), probability forecasts from the statistical downscaling method applied to the control member of the EPS (red) and probability forecasts from the statistical downscaling method applied to the whole EPS (green). The horizontal black line represents the no-skill line.


Brier Skill Score v Lead Time (Hours)
ALL 5 STATIONS – Pr{Y>0.1mm}
EPS+MIX 2003
EPS 2003
CM+MIX 2003
Brier Skill Score
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
0
24
48
72
96
120
144
168
192
216
240
Forecast Lead time


Brier Skill Score v Lead Time (Hours)
ALL 5 STATIONS – Pr{Y>1.0mm}
EPS+MIX 2003
EPS 2003
CM+MIX 2003
Brier Skill Score
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
0
24
48
72
96
120
144
168
192
216
240
Forecast Lead time

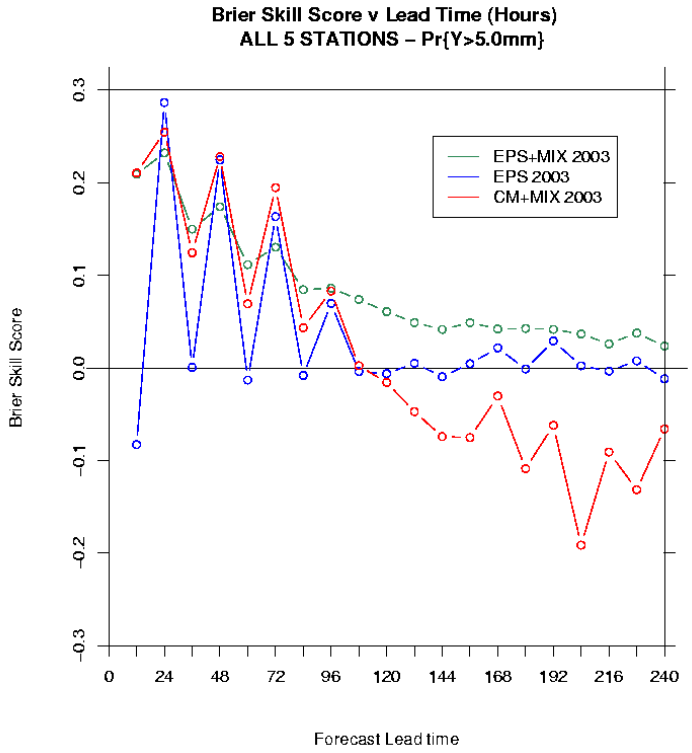


Figure 2, Plots of BSS against forecast lead time in hours for the a) 0.1mm, b) 1.0mm and c) 5.0mm exceedence thresholds. For probability forecasts from: the raw ECMWF EPS (blue), the statistical downscaling method applied to the control member of the EPS (red) and the statistical downscaling method applied to the whole EPS (green).

In this study the dry-wet threshold was chosen to be 0.1mm. Events where precipitation greater than 0.1mm is observed / forecast are defined to be wet; events where precipitation less than or equal to 0.1mm are observed / forecast are defined to be dry. Figure 2a shows that at the 0.1mm exceedence threshold the raw EPS performs worse than the simple climatological forecast for all forecast lead times from T+12 to T+240. This is believed to be attributable to a known feature of the EPS to over-predict probabilities of small precipitation accumulations (Mylne, personal communication). There is also a distinct diurnal oscillation in the BSS obtained at 00UTC and 12UTC; the forecasts for the twelve hours preceding midnight consistently achieve a better BSS than the EPS forecasts for the twelve hours preceding midday. There are a number of possible explanations for this behaviour including a coding error, sampling variations between 00UTC and 12UTC and the behaviour of the ECMWF model. The code used to generate the statistics has been checked a number of times and there is no evidence of any coding errors; we therefore proceed on the assumption that the code is correct. If the behaviour is a result of sampling variations at 00UTC and 12UTC one would also expect an oscillation of a similar magnitude to be present in the Brier Score results for the simple climatological forecast, however, this is not observed (not shown). Therefore, it is concluded that the oscillation in the BSS is more likely to be attributable to the behaviour of the ECMWF system. Two possible explanations for this behaviour are proposed:

- Overnight the ECMWF model is subject to radiative cooling resulting in the formation of low stratus cloud. If excessive low stratus cloud is formed there will be a corresponding tendency of the model to produce light rain in the twelve hours preceding midday.

- It might be expected that the convection scheme will be triggered more often in the afternoon than in the morning. If the convection scheme triggers precipitation it will rain excess moisture out from the model and hence produce less light drizzle than in the morning.

Figure 2a also shows that when the statistical downscaling method is applied to the control member of the EPS (red), the BSS remains positive compared to the simple climatological forecast until day six (T+144). This is a striking result that a single model forecast plus a statistical correction can beat the 51-member EPS. However, this must be seen in context - the EPS is affected by a small bias which happens to be critical for the 0.1mm threshold when measured by the BSS and which is easily corrected by the downscaling.

The statistical model applied to the control member of the EPS makes no estimate of synoptic scale uncertainty and cannot account for random poor model forecasts. Therefore, the statistical downscaling method was applied to the whole EPS; Figure 2a shows that for these probability forecasts the BSS remains positive at all lead times and improves upon the statistical model applied to the control member of the EPS as well as the raw EPS at all lead times. The benefit of applying the statistical downscaling to the whole EPS is particularly apparent after day five. The choice of the wet-dry threshold of 0.1mm in the development of the statistical model implicitly accounts for the systematic tendency of the ECMWF model to over-predict small precipitation accumulations and effectively acts as a bias correction.

Figure 2b presents a similar plot of BSS versus forecast lead time for the 1.0mm threshold. The results show that the statistical model applied to the whole EPS (green) performs better than the raw EPS (blue) and retains a positive BSS at all forecast lead times. Figure 2b also indicates that until approximately day four, the forecasts obtained from the statistical model applied to the control member of the EPS (red) perform similarly in terms of the BSS to forecasts obtained from the statistical model applied to the whole EPS. However, after day four the skill of the statistical model applied to the control member of the EPS drops sharply and becomes negative compared to the simple climatological forecast at day six, reflecting the inability of the forecast to estimate the synoptic scale uncertainty.

Figure 2c presents a plot of BSS versus forecast lead time for the 5.0mm threshold. The results show that prior to day four (T+96) the BSS for the raw EPS exhibits a significant diurnal variation; forecasts for the twelve hours preceding midnight do not improve upon the simple climatological forecast during this period. It should be noted that this variation in the behaviour of the raw EPS is in the opposite sense to that observed at the 0.1mm threshold,

Figure 2a. A possible explanation for this behaviour is that during the twelve hours preceding mid-day precipitation is dominated by the well-resolved dynamical precipitation. Precipitation during the twelve hours preceding midnight (afternoon/evening) is more strongly affected by unresolved convection. After day four the skill of the raw EPS is marginal compared to the basic climatological forecast. Figure 2c shows that probability forecasts obtained through the application of the statistical model to the control member of the EPS perform better than probability forecasts from the raw EPS until day five, after which the skill becomes negative. The statistical model applied to the whole EPS generates a large improvement in the Brier skill score for forecasts verifying at 00UTC at the short forecast ranges. This would be expected if the events generating more than 5mm of precipitation, in the twelve hours preceding 00UTC, are dominated by processes that operate on scales unresolved by the EPS. The skill of the forecasts generated by applying the statistical model to the whole EPS, compared to the simple climatological forecast, remains positive at all time ranges.

An important consideration when analysing the results is whether the null hypothesis, "the results have arisen by chance", can be rejected. Traditional significance tests are inappropriate in this case and bootstrap re-sampling (Wilks, 1995) was used to generate confidence intervals around the Brier Score. Figure 3 presents plots of the Brier Score (solid lines) and the corresponding 95% confidence intervals (dashed lines) versus forecast lead time for the 0.1mm and 5.0mm thresholds. Only the results for the raw EPS (blue) and the statistical model applied to the whole EPS (green) are presented here. Figure 3a shows that at the 0.1mm threshold the improvement observed over the EPS is statistically significant at the 0.05 level. At the higher thresholds, Figure 3b, the improvement is not statistically significant. Whilst it is not clear that the significance testing results found for the Brier score can be extended to the BSS it seems unlikely that the results would change greatly. It should be noted that the strong diurnal variation observed in the BSS for the raw EPS probability forecasts at the 5.0mm exceedence threshold, Figure 2c, is not apparent in Figure 3b. However, a diurnal variation is observed in the climatological forecast (not shown) which translates, through equation 2, to the pronounced diurnal variation observed in the BSS in Figure 2c. The diurnal variation observed in the climatological forecast occurs because separate climatological forecasts (based on past observations taken at 00UTC and 12UTC) were issued for forecasts verifying at 00UTC and 12UTC.

Summarising, the statistical model applied to the whole EPS retains a positive BSS when compared to a simple climatological forecast at all forecast lead times and at all exceedence thresholds. However, the magnitude of the improvement in skill decreases with increasing exceedence threshold. The skill exhibited by the statistical model applied to the whole EPS is also superior to the skill exhibited by the raw EPS at all forecast lead times and thresholds. The benefit of applying the statistical model to the whole EPS

and not just to the control member of the EPS is particularly evident at the longest forecast lead times. Unfortunately, results of the statistical significance testing applied to the Brier Score indicate that the improvement observed in the Brier score is only significant at the 0.1mm threshold.

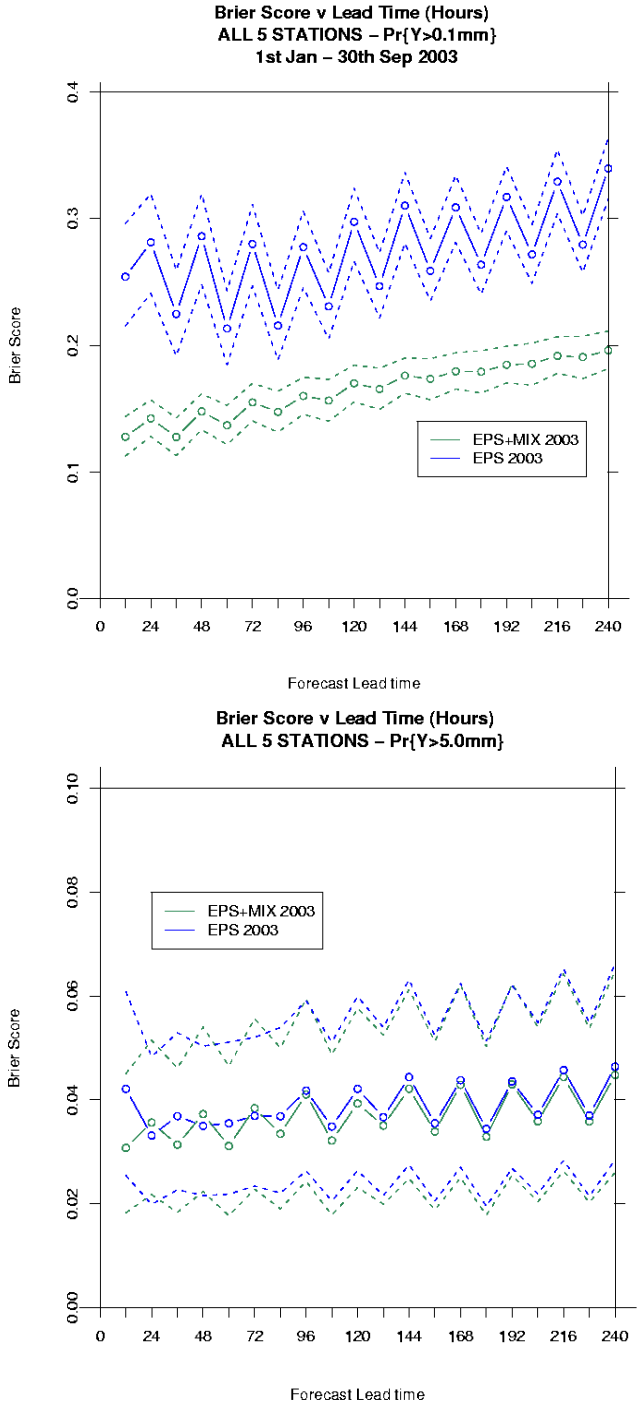


Figure 3, Plots of Brier Score (solid) and the 95% Confidence intervals (dashed) against forecast lead time in hours for a) 0.1mm, b) 5.0mm. For probability forecasts from: the raw ECMWF EPS (blue), the statistical downscaling method applied to the control member of the EPS (red) and the statistical downscaling method applied to the whole EPS (green).

b) Reliability

Murphy (1973) described a partition of the Brier Score into three component parts: reliability, resolution and uncertainty. The reliability of a forecast system is a measure of its ability to provide an unbiased estimate of the observed frequency of an event associated with different probability values. The reliability diagram, where the observed relative frequency of an event is plotted against the forecast probability (see Figure 4), provides a way to visually assess the degree of reliability.

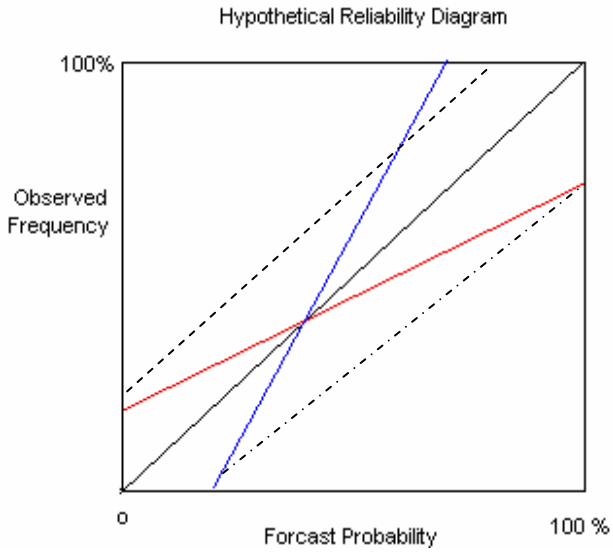


Figure 4. A hypothetical Reliability Diagram, the black diagonal line represents a forecast system that is perfectly reliable. The dashed line underneath (above) the diagonal represents a forecast system that systematically over-forecasts (under-forecasts) probabilities. The red line represents an over-confident forecast system, where low probabilities are under-forecast and high probabilities are over-forecast. Conversely the blue line represents an under-confident forecast system.

The reliability diagram is often presented in association with the sharpness diagram which summarises the frequency of use of each probability forecast value. Reliability and sharpness diagrams are presented for probability forecasts obtained from the raw EPS (blue), the statistical model applied to the control member of the EPS (red) and the statistical model applied to the whole EPS (green). A perfectly reliable forecast will lie along the diagonal, where the observed relative frequency and the forecast probability are equal. Deviations from the diagonal usually represent forecast biases as explained in Figure 4. However, in a small sample deviations from the diagonal can be a result of sampling variations, limiting the usefulness of the diagnostic in assessing the forecast quality (Wilks, 1995). Sampling variations are especially evident in the verification results for the larger exceedence thresholds, see Figure 5a. Despite the obvious sampling variations in Figure 5a, it can be seen that the statistical downscaling method eliminates the high probabilities forecast by the raw EPS (as shown in the sharpness diagram) which are not justified by the reliability diagram. It is also evident from Figure 5a that there is only a small difference in the performance of probability

forecasts generated by the statistical model applied to the control member of the EPS (red) and whole EPS (green). This result indicates that, as expected, the spread of ensemble is small at the T+12 forecast lead time; the ECMWF ensemble is designed to estimate the uncertainty at the 3-10 day forecast range. Figure 5b shows that the reliability of the probability forecasts is improved through the application of the statistical model to the whole EPS compared to just the control member of the EPS at T+108 (day 4.5) when synoptic-scale uncertainty is important and the spread of the ECMWF EPS is larger.

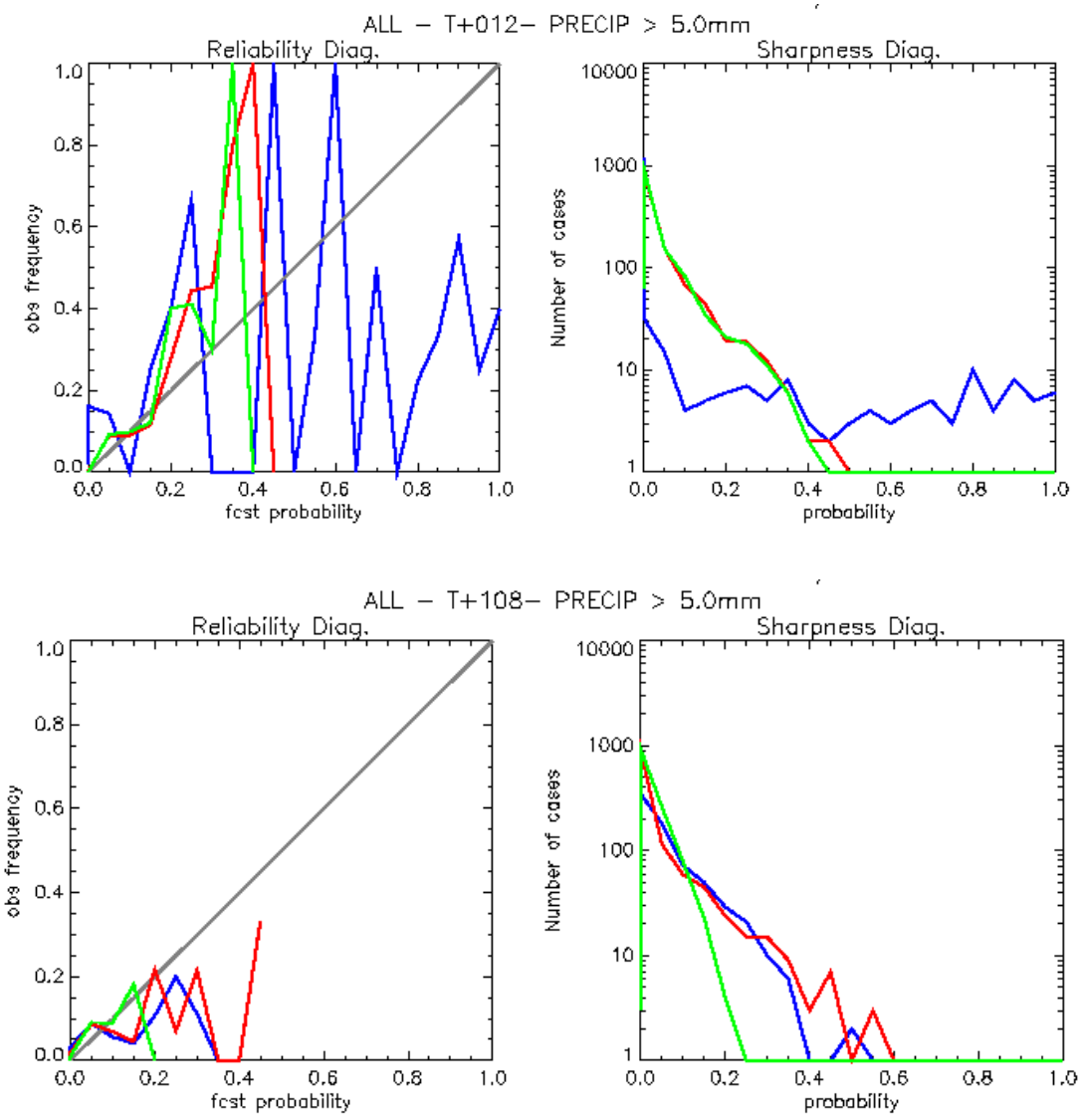


Figure 5. Reliability and sharpness diagrams for the raw EPS (blue), the statistical model applied to the control member of the EPS (red) and the statistical model applied to the whole EPS (green). The results are for the 5.0mm threshold at a) T+12 and b) T+108.

At the 0.1mm threshold the probability forecasts from the raw EPS are observed to be unreliable, systematically over-predicting the probability of a wet event at all forecast lead times. This is corrected through the application of the statistical downscaling method to the EPS to produce more reliable

forecasts, Figure 6. The results presented in Figure 6 for the statistical model applied to the control member of the EPS (red line) indicate that at the 0.1mm (dry-wet) threshold a limited range of probabilities are forecast for precipitation exceeding 0.1mm. This may appear strange at first; one might expect that a larger range of probabilities should be predicted as seen in Figure 5 for the 5mm threshold. However, Equations A.16 – A.18 indicate that the probability of precipitation exceeding the 0.1mm threshold reduces to a fixed estimate based on the values of π_{dd}, π_{dw}, π_{wd} and π_{ww}. Therefore, the range of probabilities at the 0.1mm threshold predicted is limited when averaged over the five sites used in this study.

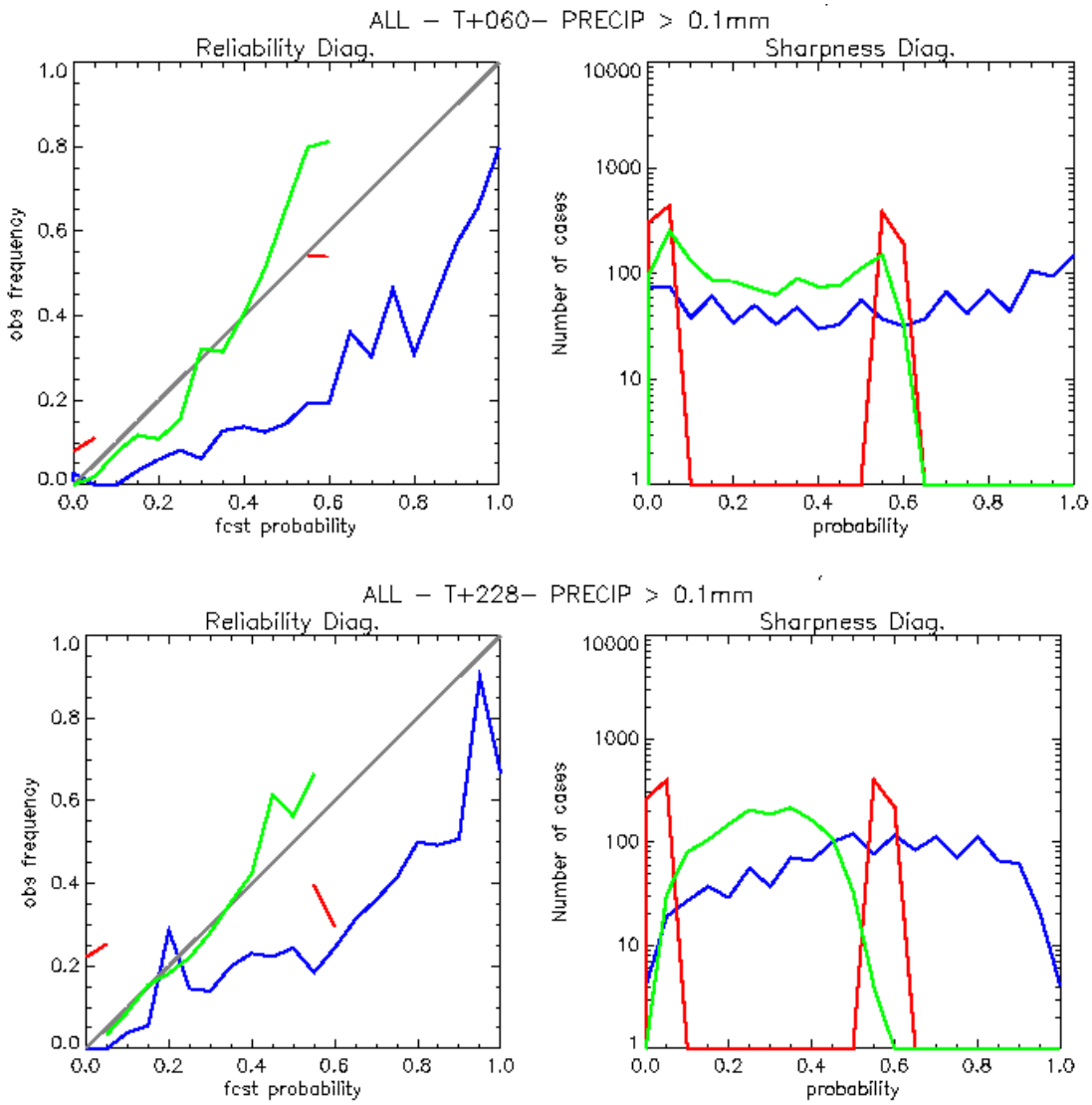


Figure 6. Reliability and sharpness diagrams for the raw EPS (blue), the statistical model applied to the control member of the EPS (red) and the statistical model applied to the whole EPS (green). The results are for the 0.1mm (dry/wet) threshold at T+60 (top) and T+228 (bottom)

The results for the statistical model applied to the EPS at higher thresholds, above 5.0mm (not shown), are typical of well-calibrated forecasts of rare events. At low forecast probabilities the reliability curve lies very close to the

diagonal; at the higher probabilities there are deviations away from the diagonal but these are subject to small samples. Although the forecasts show good reliability they possess little resolution because only very low probabilities are ever issued.

c) Relative Operating Characteristic (ROC)

ROC assesses the ability of forecasts to aid decision making by measuring the strength of the signal in the forecast which differentiates occurrences of weather events from non–occurrences. For a probabilistic forecast the ROC curve is a plot of hit rate versus false alarm rate at various decision (or probability) thresholds, Figure 7. The area underneath the ROC curve is a measure of the intrinsic capability of the forecasting system to discriminate between events and non-events. Theory determines that a climatological forecast which has no discriminatory skill will possess a ROC that is the diagonal passing through the points (0,0) and (1,1); the area under the no-skill line is 0.5. Therefore, for a forecast system to possess skill, A should be greater than 0.5 and the larger the value of A the better the discriminatory ability of the forecast system.

There are two different methods of generating a ROC curve, one is parametric and the other is empirical (Mason, 2003). For the purposes of this report the empirical ROC was plotted by simply joining the data points with straight lines. However, as pointed out by Wilson (2000) this is a crude method of fitting a ROC and A will be underestimated using this method compared to the un-biased value of A derived from the parametrically-fitted ROC.

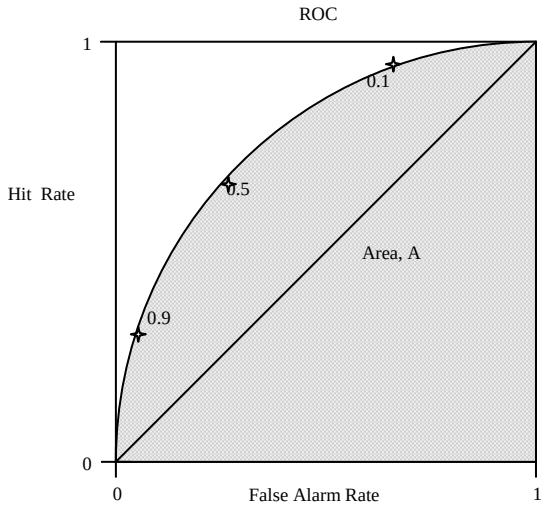
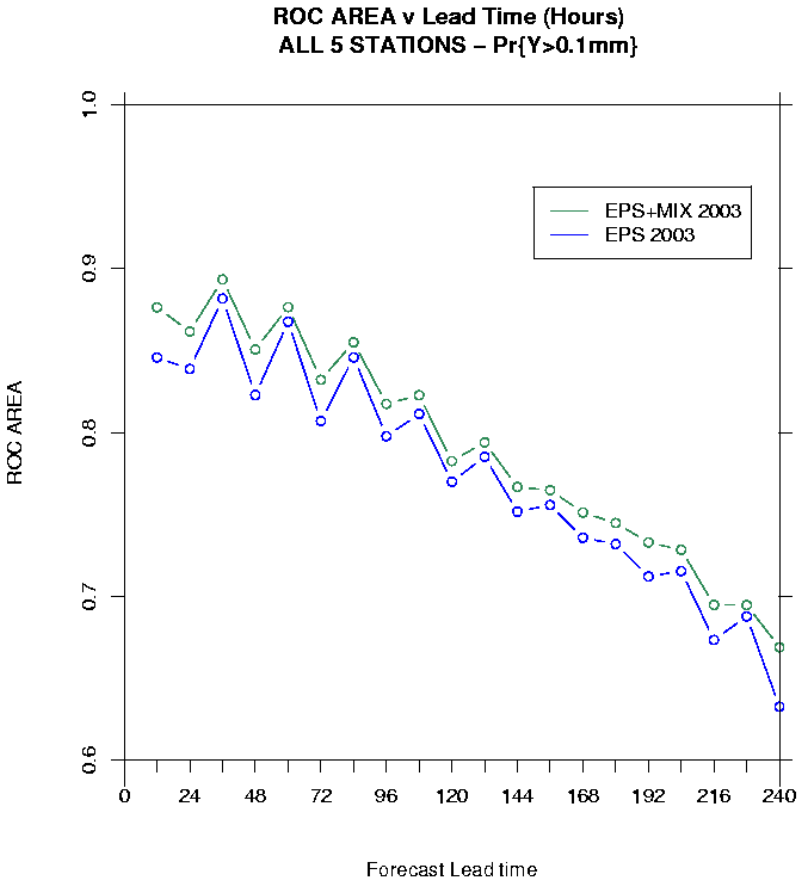
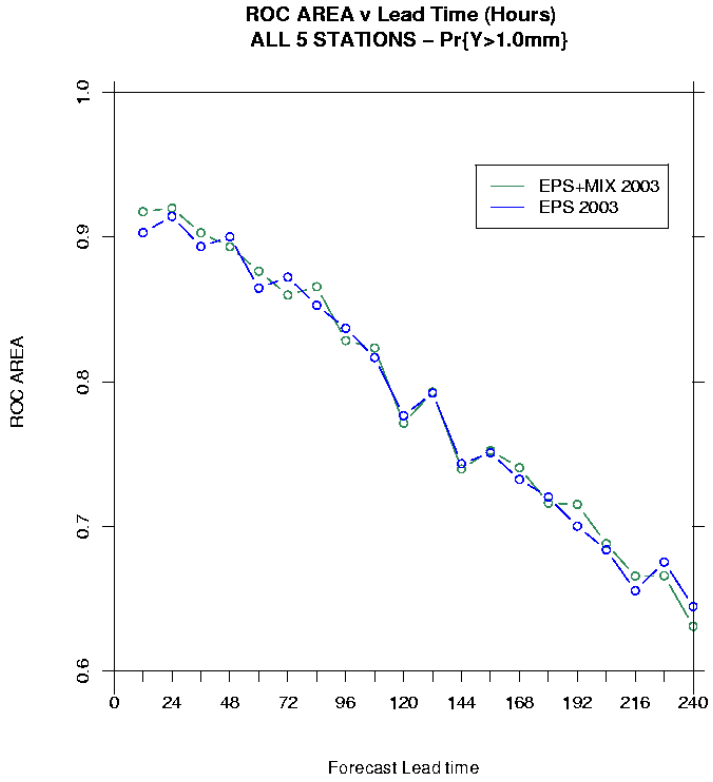


Figure 7. An example ROC curve, the values 0.1, 0.5 and 0.9 correspond to various probability (decision) thresholds. The shaded area corresponds to the area under the curve, A. The diagonal line passing through (0,0) and (1,1) represents the no-skill line.

Figure 8.0 presents plots of A, the area under the ROC curve, against forecast lead time for all stations at the precipitation thresholds 0.1mm, 1.0mm and 5.0mm respectively. Each plot shows the area under the ROC curves for probabilistic forecasts obtained from the raw EPS data (blue) and for forecasts obtained from the statistical downscaling method applied to the whole EPS (green). Figure 8a shows that the area under the ROC curve, A, decreases with increasing forecast lead time for both forecast systems, but that it remains above the no-skill threshold of 0.5 at all forecast lead times. It is clear that at the 0.1mm threshold (which corresponds to the dry-wet threshold) applying the statistical model to the EPS improves the ability of the system to discriminate between wet and dry events at all forecast lead times. This is evidence that the statistical downscaling method is also providing a bias correction, implicitly accounting for the tendency of NWP models to over-predict very small precipitation accumulations, through the separation of the problem into wet and dry event distributions. At the 1.0mm threshold both systems perform similarly and the benefit observed at the 0.1mm threshold is not apparent. At the 5.0mm threshold the statistical downscaling applied to the EPS improves upon the raw EPS for the first 48 hours after which the performance is similar to the raw EPS.





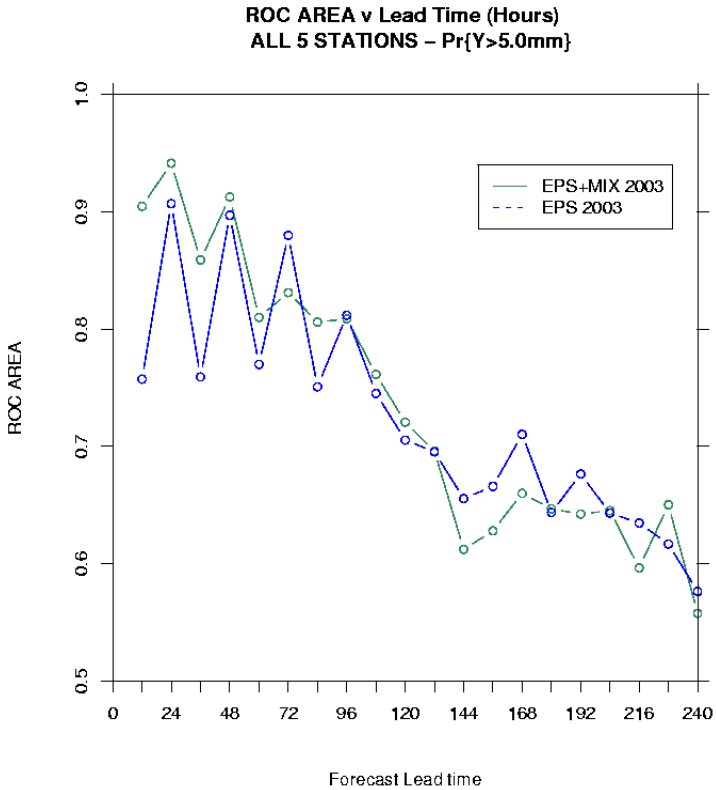


Figure 8. Area under ROC curve against forecast lead time for the raw EPS (blue) and the statistical model applied to the whole EPS (green) at a) 0.1mm, b) 1.0mm and c) 5.0mm.

d) Economic Value

Forecast users often require a forecast to help them choose between different courses of action and are usually more interested in the economic value of the forecast than in scientifically orientated verification measures. Each action taken by a user possesses an associated cost and leads to an economic benefit/loss depending on whether a particular weather event occurs. Given information regarding costs incurred by the user when protective action is taken and expected losses when protective action is not taken and an adverse weather event occurs, a cost/loss ratio can be defined for each user. The economic value of a forecast can be defined as follows:

$$Value = \frac{E_{climate} - E_{forecast}}{E_{climate} - E_{perfect}} \quad . \qquad (3)$$

$E_{climate}$ is the mean expense incurred by the user if a climatological forecast is used, $E_{perfect}$ is the mean expense for a perfect forecast and $E_{forecast}$ is the mean expense for the forecast system under investigation (Richardson, 2003). For a given forecast system it is possible to calculate the economic value expected for a range of cost/loss ratios; see Figure 9. Figure 9 presents plots of the economic value for the raw EPS (blue), the statistical model applied to the control member of the EPS (red) and the statistical model applied to the whole EPS (green).

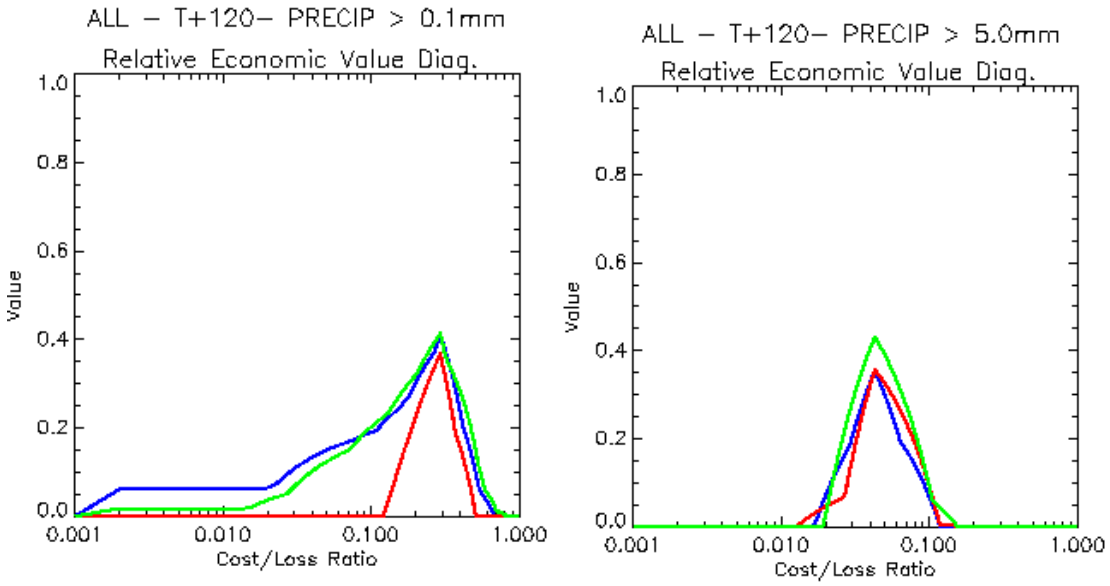


Figure 9. Relative economic value versus the cost/loss ratio at T+120 for a) the 0.1mm exceedence threshold (left) and b) the 5.0mm threshold (right). Results are presented for the raw EPS (blue), the statistical model applied to the control member of the EPS (red) and the statistical model applied to the whole EPS (green).

A number of general conclusions can be drawn from the verification results for the economic value. For all three types of probability forecast, at each

forecast lead time, the position of the maximum value, $V_{max,}$ moves to smaller cost/loss ratios as the exceedence threshold increases and the events become rarer. This is illustrated for day five forecasts (T+120) in Figure 9. Similarly, Figure 9 shows that the range of cost/loss ratios with positive economic value decreases with increasing exceedence threshold. The results presented in Figure 10 show that as the forecast lead time increases the magnitude of V_{max} decreases at the 1.0mm exceedence threshold. This result is consistent with the observed drop in the BSS and the area under the ROC curves at longer forecast lead times and is valid at all thresholds (not shown).

The results also indicate that at the shorter forecast lead times (T+12 to T+96) the range of cost/loss ratios spanned by the economic value curve for each forecast threshold is increased by applying the statistical downscaling method to the whole EPS compared to just using the raw EPS, Figure 10a. The benefit of applying the statistical downscaling to the EPS at short time ranges is particularly apparent at the smallest cost/loss ratios; these forecasts are beneficial to users who are sensitive to rare events. After day 2 (T+48) it is clear that the forecasts obtained from the raw EPS and the statistical model applied to the whole EPS perform better in terms of economic value than forecasts obtained from the statistical model applied to only the control member of the EPS, Figure 10b. This is also consistent with the previously presented results and confirms that after day 2 the EPS provides a valuable sampling of the synoptic-scale uncertainty not provided by the statistical model alone. It should be noted that this result is clearly evident in the economic value results for thresholds less than 5mm, however, it is not obvious at the larger thresholds (see Figure 9b).

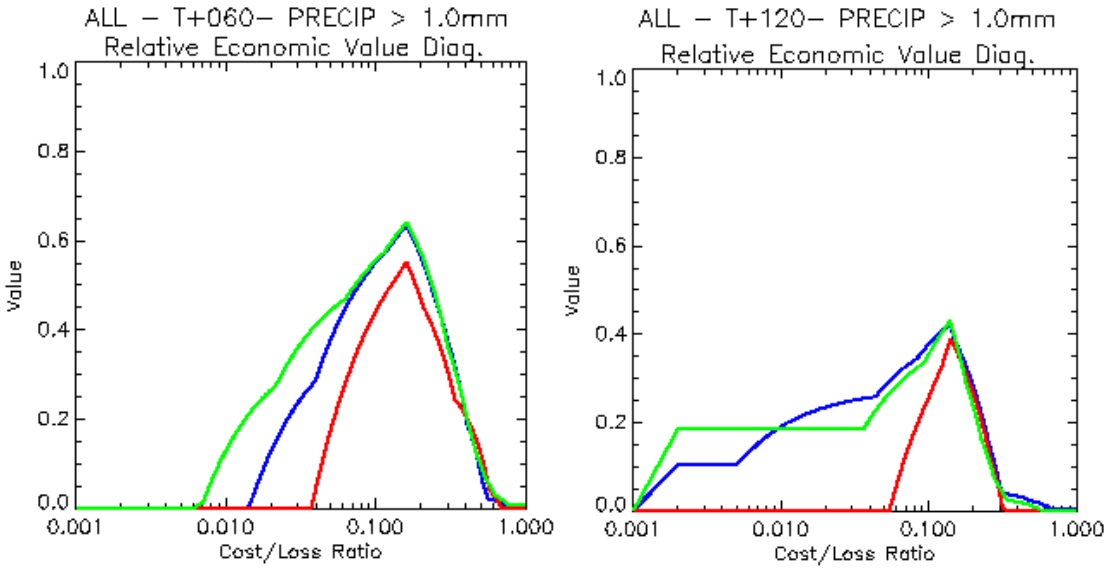


Figure 10. Relative economic value versus the cost/loss ratio for the 1.0mm exceedence threshold at a) T+060 (left) and b) at T+120 (right). Results are presented for the raw EPS (blue), the statistical model applied to the control member of the EPS (red) and the statistical model applied to the whole EPS (green).

At the longest forecast lead times (days 9.5 and 10) there is no obvious benefit in using the statistical model applied to the EPS (not shown); the magnitude of V_{max} is similar for both the raw EPS and the statistical model applied to the whole EPS.

e) Sensitivity experiment

Precipitation data are highly skewed. Therefore, prior to calculating the model parameters, the training data is transformed to become more normally distributed, using the Box-Cox transformation (Appendix A). The degree of transformation applied by the Box-Cox method is determined by the value of the tuneable parameter lambda. The tuning of lambda was initially carried out subjectively viewing quantile plots and histograms of the data. This resulted in different values of lambda, in the range 1/3 to 1/5, being assigned to each site. In order to keep up-to-date with changes in the ECMWF model, the statistical model parameters will need to be updated as the training data is updated. It is not clear that the value of lambda will remain constant as the training data changes and an automatic re-selection process to determine the optimum value of lambda would be desirable. However, it is not obvious how to automate the selection of lambda; therefore, an experiment to consider the sensitivity of the BSS to the value of lambda selected was performed. Values of lambda in the range of one half to one sixth were selected for testing based on the subjective assessment of quantile and histogram plots of the data.

Figure 11 is a plot of the BSS against forecast lead time for the 1.0mm exceedence threshold at Manchester Ringway. The results presented are for the raw EPS and the statistical model applied to the EPS with lambdas in the range one half to one sixth. The overall result of the sensitivity test for the five selected sites indicates that the value of lambda equal to one half (red line in Figure 11) clearly produces worse results in terms of the BSS when compared to values in the range one third to one sixth. There is some evidence (not shown) that the value of lambda equal to one third performs less well than lambda's in the range one fourth to one sixth for all sites but the difference is small compared to that between one third and one half. Therefore, we conclude that the BSS results are relatively insensitive to the value of lambda in the range one third to one sixth. It is recommended that in future work a value of 1/4 should be used for lambda for all sites.

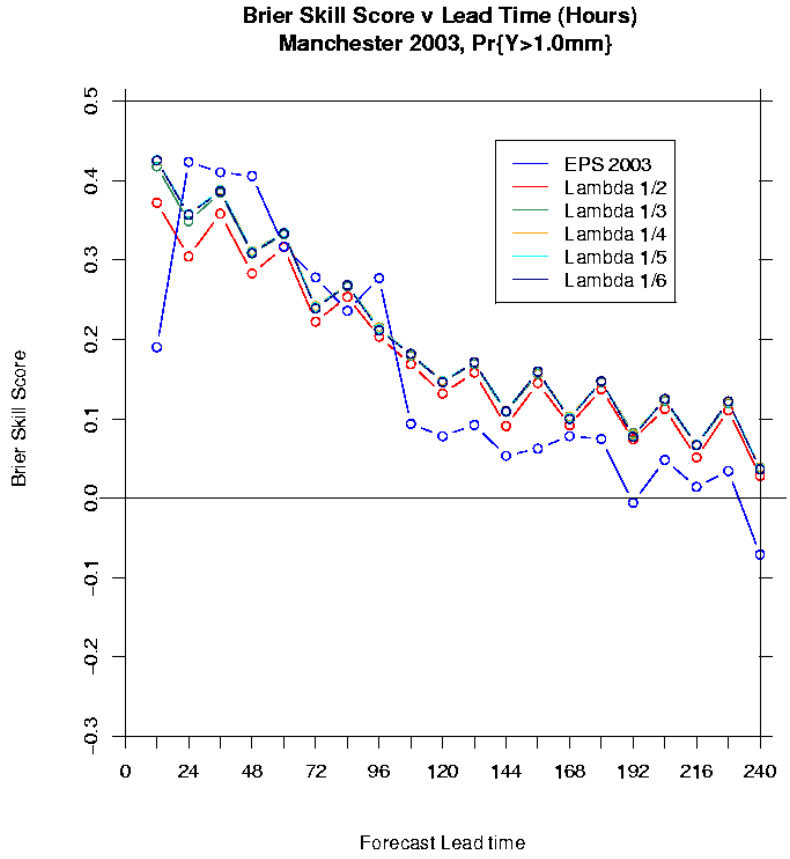


Figure 11, Brier Skill Score against forecast lead time for Manchester (03334) for the raw EPS and the statistical model applied to the EPS with lambdas in the range one half to one sixth

4. Discussion and Conclusions

The main purpose of this work has been to develop a method capable of producing reliable site-specific probability of precipitation forecasts from the ECMWF EPS using statistical downscaling. A further aim has been to assess whether any benefit realised from using the statistical downscaling justifies including this method in the operational PREVIN system run at the Met Office. The method has been developed and tested for 5 sites for the period 1st January – 30th September 2003. The results from these tests indicate, as expected, that the statistical downscaling does improve the reliability of the probabilistic forecasts from the raw EPS. However, the ROC results indicate that the discrimination is not improved (except at the 0.1mm threshold). The results also show that the benefit of applying the downscaling to the whole EPS compared to just the control member of the EPS is particularly evident at the longer forecast lead times. This justifies the original assumption that applying the statistical model to the EPS would account for the uncertainty due to both the unresolved and synoptic scales. The statistical downscaling applied to the whole EPS exhibits the most skill in terms of the Brier Score, ROC area and reliability at the 0.1mm (dry/wet threshold). The statistical model used implicitly accounts for the tendency of NWP models to over-

predict very small precipitation accumulations. At the larger thresholds the Brier Skill score for the statistical downscaling applied to the whole EPS remains positive and larger than the Brier Skill score of the raw EPS at all forecast time ranges. However, at the larger thresholds the improvements observed in the Brier Score are not statistically significant at the 0.05 level.

Whilst these results do give an indication of the performance of forecasts generated through the application of statistical downscaling to the EPS, problems with the experimental design determine that they cannot provide sufficient evidence to justify implementing the method operationally. Due to issues over data availability the forecasts were only verified over a nine-month period which means that not all seasons were represented in the verification statistics. Furthermore, 2003 was particularly dry during the spring and summer which is not ideal for verifying precipitation forecasts, especially forecasts for large precipitation accumulations.

The most important consideration must be the forecast system chosen to compare the downscaled probability forecasts against. In this study raw probability forecasts from the ECMWF ensemble prediction system were chosen. However, retrospectively it is clear that to obtain a fair assessment of the potential benefit of the new system it should have been compared to the performance of the current calibration system operational in PREVIN. Therefore, it is recommended that the statistical downscaling method should be assessed over an extended period of time against the current operational system used in PREVIN; the longer the time series used for verification purposes the more confidence can be placed in the results.

5. Acknowledgements

I would like to thank David Stephenson at the University of Reading Meteorology Department for all his invaluable help in the development of this statistical model. My thanks are also due to Caroline Woolcock and Tim Legg at the Met Office for providing me with the data used in this study.

Appendix A

Time series of precipitation observations indicate that precipitation is characterised by a mixture of wet and dry events: rain is either observed at an individual site or it is not. This property of precipitation was recognised by Katz (1977) who described the distribution of daily precipitation accumulations in terms of the separation between wet and dry events. Wilks (1990) and Krzysztofowicz and Sigrest (1997) do consider the wet/dry distinction in terms of generating local climatic guidance based on long historical observation records; however, very limited reference has been made in the previous literature to the distinction between the wet and dry event PDFs with regard to site-specific forecasts of precipitation amount. Thus, following the approach of Katz (1977), the probability density function

(PDF) of accumulated precipitation $f(y)$, can therefore be modelled as a mixture of the wet and dry event PDFs, $f_w(y, \varepsilon)$ and $f_d(y, \varepsilon)$ respectively:

$$f(y) = f_d(y, \varepsilon)\pi_d(\varepsilon) + f_w(y, \varepsilon)\pi_w(\varepsilon) \quad \text{(A.0)}$$

$$\int_u^\infty f_w(y, \varepsilon)dy = 1 \quad \text{(A.1)}$$

$$\int_o^u f_d(y, \varepsilon)dy = 1 \quad \text{(A.2)}$$

where ε is the threshold delineating a wet from a dry event, $\pi_w(\varepsilon)$ represents the probability of a wet event as a function of the threshold defining measurable precipitation, ε, and $\pi_d(\varepsilon) = 1 - \pi_w(\varepsilon)$ is the probability of a dry event, which is simply the complement of $\pi_w(\varepsilon)$. As the threshold defining a wet event decreases towards zero it converges to a constant π_w and in the limit as ε decreases towards zero, $f(y)$ becomes

$$Y \sim \lim_{u \to 0} f(y) \to \delta(y)\pi_d + f_w(y).\pi_w \quad \text{(A.3)}$$

where $f_w(y)$ is the wet event PDF, $\delta(y)$ is the delta function, which the dry event PDF is reduced to in the limit. Hence, when the probability of accumulated precipitation is calculated it should be treated as combination of the wet event PDF and the dry event PDF. Similarly, the forecasts of precipitation, X, can be modelled as the mixture distribution:

$$X \sim f(x) = \delta(x)\pi_d + f_w(x).\pi_w \quad \text{(A.4)}$$

where $f_w(x)$ is the wet event PDF and $\delta(x)$ is the delta function representing the dry event PDF, whilst π_d and π_w are the probabilities of dry and wet events respectively. In an extension of Eqn's (A.3 and A.4), bivariate precipitation data, consisting of corresponding forecasts, X, and observations, Y, can be modelled as a four-component mixture model of wet and dry event PDFs (Stephenson, 2003, personal communication)[2]:

$$f(x, y) = \pi_{dd}\delta(x)\delta(y) + \pi_{dw}\delta(x)\, f_y(y) + \pi_{wd}\delta(y)\, f_x(x) + \pi_{ww}h(x, y) \quad \text{(A.5)}$$

where

$\pi_{ww} = \Pr\{ X > \varepsilon \text{ and } Y > \varepsilon \}.$

$\pi_{wd} = Pr\{ X > \varepsilon \text{ and } Y \leq \varepsilon \}.$

$\pi_{dw} = Pr\{ X \leq \varepsilon \text{ and } Y > \varepsilon \}$

[2] Equation A.5 provides a full description of the problem and not all the distributions contained in the equation are required to make a probability forecast; however, the extra information is included here for completeness.

$\pi_{dd} = Pr\{ X \leq \varepsilon \text{ and } Y \leq \varepsilon \}$,

$\pi_{dd} + \pi_{dw} + \pi_{wd} + \pi_{ww} = 1$

and

$$X \mid (Y \leq \varepsilon \text{ and } X > \varepsilon) \sim f_x(x)$$
$$Y \mid (X \leq \varepsilon \text{ and } Y > \varepsilon) \sim f_y(y)$$
$$X,Y \mid (Y > \varepsilon \text{ and } X > \varepsilon) \sim h(x,y)$$

π_{dd} is the probability that both the forecast and the corresponding observations are dry, a dry-dry event; π_{dw} is the probability that the forecast is dry and the observation is wet, a dry-wet event; π_{wd} is the probability that the forecast is wet and the observation is dry, a wet-dry event; π_{ww} is the probability that both the forecasts and the observations are wet, a wet-wet event. $\delta(x)$ and $\delta(y)$ are delta functions representing the distributions of X and Y for dry-dry events. $f_x(x)$ is the marginal distribution of the forecasts for wet-dry events and $f_y(y)$ is the marginal distribution of the observations for dry-wet events. Finally, $h(x,y)$ is the joint distribution of precipitation for wet-wet events. Eqn (A.5) is illustrated schematically in Figure 1.0.

The probabilities π_{dd}, π_{dw}, π_{wd}, and π_{ww} can be estimated from the proportion of the total observations, n, that represent dry-dry, dry-wet, wet-dry and wet-wet events respectively (the frequencies in a two by two contingency table of wet and dry events for the forecasts and observations). The PDFs $f_x(x)$, $f_y(y)$ and $h(x,y)$ can be estimated by fitting theoretical probability distributions to the corresponding forecasts and observations of precipitation, for wet-dry, dry-wet and wet-wet events respectively.

Precipitation accumulation data is highly skewed, with a long tail extending to large positive values. However, the data can be transformed to become more normally distributed - prior to fitting the PDFs $f_x(x)$, $f_y(y)$ and $h(x,y)$ - using the Box-Cox transformation:

$$g(x-\varepsilon) = \frac{(x-\varepsilon)^{\lambda} - 1}{\lambda} \qquad \lambda \neq 0 \qquad \text{(A.6)}$$

where λ is the transformation parameter which can be tuned appropriately to generate normally distributed data, x is either the forecast X or the observation Y, and ε is the threshold delineating a wet from a dry day. The distributions of the transformed precipitation (X',Y') can then be modelled as normal distributions with two parameters: the mean and the standard deviation.

$$X' \sim f_x(x) = N(\mu_{X'}, \sigma^2_{X'}) \qquad \text{(A.7)}$$
$$Y' \sim f_y(y) = N(\mu_{Y'}, \sigma^2_{Y'}) \qquad \text{(A.8)}$$

The distribution of precipitation for wet-wet events, *h(x,y)*, can be described by a bivariate normal distribution with five parameters. For wet-wet events the transformed forecasts, $X"$, and observations, $Y"$, can each be modelled by a normal distribution, with the associated mean and standard deviation:

$$X" \sim N(\mu_{X'}, \sigma^2_{X'}) \quad \text{(A.9)}$$

$$Y" \sim N(\mu_{Y'}, \sigma^2_{Y'}) \quad \text{(A.10)}$$

These two distributions are also related by a correlation coefficient, ρ. If the forecasts and observations are normally distributed a linear regression fit can also be applied to the data,

$$Y" = \alpha_0 + \alpha_1 X" + \sigma^2_\varepsilon \quad \text{(A.11)}$$

which contains three parameters the coefficients α_0 and α_1 and the standard error on the fit, σ_ε. Linear least squares regression is based on the assumption that the residuals are random variables scattered about a mean value (the regression line) with constant variance. The assumption that the residuals are identically and normally distributed can be used to help choose the distribution of the transformed observations, $Y"$, given the forecast, $X"$.

$$Y" \big| X" \sim N(\alpha_0 + \alpha_1 X", \sigma^2_\varepsilon) \quad \text{(A.12)}$$

Therefore, the coefficients α_0 and α_1 and the standard error on the fit, σ_ε, can be related to the mean and standard deviation of $Y"$ and $X"$.

$$\alpha_0 = \mu_{Y"} - \mu_{X"} \quad \text{(A.13)}$$

$$\alpha_1 = \rho \frac{\sigma_{Y"}}{\sigma_{X"}} \quad \text{(A.14)}$$

$$\sigma_\varepsilon = \sigma_{Y"} \sqrt{1 - \rho^2} \quad \text{(A.15)}$$

In summary, the bivariate mixture model, for corresponding precipitation forecasts and observations, can be characterised by the thirteen parameters listed in Table A1. The nine highlighted parameters are required to form a probability prediction, given a forecast X, the values used for each station used in this study are provided in Tables A2 and A3 for 00 UTC and 12UTC respectively.

Parameters	Estimation Method
$\pi_{dd}, \pi_{dw}, \pi_{wd}, \pi_{ww},$	Estimate from contingency table of event precipitation for X and Y
$\mu_{X'}, \sigma^2_{X'}$ for $f_x(x)$	Estimate by calculating the mean and standard deviation of X' for wet-dry events
$\mu_{Y'}, \sigma^2_{Y'}$ for $f_y(y)$	Estimate by calculating the mean and standard deviation of Y' for dry-wet events.
$\alpha_0, \alpha_1, \sigma^2_\varepsilon$ $\mu_{X''}, \sigma^2_{X''}$	Estimate through linear regression of Y'' on X'' and calculating the mean and standard deviation of X'' for wet-wet events.

Table A1, parameters describing the bivariate mixture model of precipitation, the nine highlighted parameters are required to form a probability prediction, given a forecast X.

site	λ	π_{ww}	π_{wd}	π_{dw}	π_{dd}	$\mu_{Y'}$	$\sigma_{Y'}$	α_0	α_1	σ_ε
03334	1/5	265/671	180/671	7/671	219/671	-0.31	1.05	0.37	0.68	1.21
03772	1/4	201/672	36/672	25/672	310/672	-0.70	1.32	0.50	0.65	1.21
03063	1/5	277/646	225/646	8/646	136/646	-1.10	0.91	-0.07	0.58	1.34
03779	1/4	203/664	128/664	20/664	313/664	-0.40	1.27	0.43	0.63	1.31
03717	1/5	183/525	114/525	11/525	217/525	-0.91	1.76	0.20	0.76	1.10

Table A2. Values of the statistical model parameters, estimated at 00UTC, required to form a probability forecast for each site used in this study. The highlighted parameters correspond to the highlighted parameters in Table A1.

site	λ	π_{ww}	π_{wd}	π_{dw}	π_{dd}	$\mu_{Y'}$	$\sigma_{Y'}$	α_0	α_1	σ_ε
03334	1/5	252/680	210/680	23/680	195/680	-1.36	1.09	0.05	0.61	1.29
03772	1/3	201/680	189/680	19/680	271/680	-1.29	0.88	0.20	0.70	1.36
03063	1/5	297/651	233/651	22/651	99/651	-1.78	1.00	-0.17	0.52	1.36
03779	1/4	189/665	187/665	7/665	282/665	-1.20	0.77	0.24	0.72	1.22
03717	1/5	183/525	114/525	11/525	217/525	-0.69	1.52	0.14	0.79	1.15

Table A3. Values of the statistical model parameters, estimated at 12UTC, required to form a probability forecast for each site used in this study. The highlighted parameters correspond to the highlighted parameters in Table A1.

The definition of the threshold, ε, delineating a dry event ($X \le \varepsilon$, $Y \le \varepsilon$) from a wet event ($Y > \varepsilon$, $X > \varepsilon$) is chosen to be 0.1mm; this corresponds to the smallest exceedence threshold currently used to generate probabilities from the ECMWF EPS. The choice of a small non-zero wet/dry threshold implicitly accounts for the tendency of the ECMWF model to forecast too much light drizzle. The mixture model of corresponding forecasts and observations of precipitation can be used to generate probability forecasts of precipitation, given a single deterministic forecast value and the nine parameters highlighted in Table A1. The exceedence probability, the probability that the observed precipitation exceeds a pre-defined threshold, u, is based on the total law of probability for conditional events and can be defined as follows:

$$\Pr\{Y > u\} = 1 - \Pr\{Y \le u\} \tag{A.16}$$

$$\Pr\{Y \le u\} = \begin{cases} \dfrac{\pi_{dw} \int_{\varepsilon}^{u} f_y(y)dy + \pi_{dd}}{\pi_{dw} + \pi_{dd}} & for \quad X \le \varepsilon \\ \dfrac{\pi_{ww} \int_{\varepsilon}^{u} h(x,y)\partial y + \pi_{wd}}{\pi_{ww} + \pi_{wd}} & for \quad X > \varepsilon \end{cases} \tag{A.17}$$

If the forecast is dry, then the probability that the observation will be less than or equal to u is simply derived from $\pi_{dd,}$ π_{dw}, and the area under dry-wet distribution between ε and u. If the forecast is wet then the actual precipitation forecast amount is substituted into the regression equation, Eqn. A.11, to find the most probable estimate of the observed precipitation. Associated with the estimate of observed precipitation is an error distribution, assumed to be identically and normally distributed. The probability that the observation is less than or equal to u is simply the combination of π_{wd}, π_{ww} and the area under the error distribution between ε and u. When u=ε, the exceedence probability, Eqn. A.17, reduces to:

$$\Pr\{Y \le \varepsilon\} = \begin{cases} \dfrac{\pi_{dd}}{\pi_{dd} + \pi_{dw}} & for \quad X \le \varepsilon \\ \dfrac{\pi_{wd}}{\pi_{wd} + \pi_{ww}} & for \quad X > \varepsilon \end{cases} \tag{A.18}$$

Equation A.18 is conditional on whether the corresponding NWP model forecast is wet or dry, and represents a fixed estimate of the climatological probability of a dry event derived from the training data sample.

References

ANDERSSON, E., BELJAARS, A., BIDLOT, J., MILLER, M., SIMMONS, A., THÉPAUT,J. N., (2003). A major new cycle of the IFS: Cycle 25r4. ECMWF Newsletter No. 97 p.12-20. Spring 2003.

BRIER, G.W., (1944). 'Verification of a Forecaster's Confidence and the Use of Probability Statements in Weather Forecasting.' Research Paper No. 16, U.S. Weather Bureau, Washington, D.C.

BRIER, G. W., (1950). Verification of forecasts expressed in terms of probability. Monthly Weather Review, 78:1-3

BUIZZA, R., MILLER, M., PALMER, T.N., (1999). Stochastic representation of uncertainties in the ECMWF Ensemble Prediction System. Quarterly Journal of the Royal Meteorological Society. 125 (560): 2887-2908 Part B

BUIZZA, R., BARKMEIJER, J., PALMER, T.N., RICHARDSON D.S., (2000). Current status and future developments of the ECMWF Ensemble Prediction System. Meteorological Applications, 7 (2): 163-175

BUIZZA, R., RICHARDSON, D.S., PALMER, T.N., (2001) The 80km High-Resolution ECMWF EPS. ECMWF newsletter 90:2-9

EBERT, E.E., DAMRATH, U., WERGEN, W., BALDWIN, M.E., (2003). The WGNE Assessment of Short-Term Quantitative Precipitation Forecasts. Bulletin of the American Meteorological Society.84: 481-492

GLAHN, H.R., LOWRY, D.A.., (1972). The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. Journal of Applied Meteorology. 11:1203-1211

JOHN, S.E., (2003). Site Specific Probability of Precipitation Forecasting. MSc Dissertation. University of Reading.

KATZ, R. W., (1977). Precipitation as a chain-dependent process. Journal of Applied Meteorology 16:671-676

KRZYSZTOFOWICZ, R., SIGREST, A.A., (1997). Local Climatic Guidance for Probabilistic Quantitative Precipitation Forecasting. Monthly Weather Review, 125: 305-316

LALAURETTE, F., (2002a) Changes to the Operational Forecasting System. ECMWF Newsletter No.93 p 1-2. Spring 2002.

LALAURETTE, F., (2002b) Changes to the Operational Forecasting System. ECMWF Newsletter No.94 p 1-2. Winter 2002/2003.

LEGG, T.P., MYLNE, K.R., WOOLCOCK, C., (2002). Use of medium-range ensembles at the Met Office 1: PREVIN- a system for the production of probabilistic forecast information from the ECMWF EPS. Meteorological Applications, 9,225-271

MASON, I.B., (2003). Binary Events, in Forecast Verification: A Practioner's Guide in Atmospheric Science, Eds. I.T.Jolliffe and D.B. Stephenson, Wiley and Sons pp. 37-96

MOLTENI, F., BUIZZA, R., PALMER, T.N., PETROLIAGIS, T., (1996). The new ECMWF ensemble prediction system: methodology and validation. Quarterly Journal of the Royal Meteorological Society. 122:73-119

MURPHY, A.H., (1978). Hedging and the Mode of Expression of Weather Forecasts. Bulletin of the American Meteorological Society. 59: 371-373

MURPHY, A.H., (1973). A New Vector Partition of the Probability Score, Journal of Applied Meteorology, 12 595-600

MYLNE, K., WOOLCOCK, C., DENHOLM-PRICE, J., DARVELL, R., (2001) PREVIN - Site Specific Forecasts - User Guide. Version 1.0 - August 2001. Met Office. pp.1-30

RICHARDSON, D.S., (2003). Economic Value and Skill, in Forecast Verification: A Practioner's Guide in Atmospheric Science, Eds. I.T.Jolliffe and D.B. Stephenson, Wiley and Sons pp. 165-187

WILKS, D.S., (1990). Probabilistic Quantitative Precipitation Forecasts Derived from PoPs and Conditional Precipitation Amount Climatologies. Monthly Weather Review, 118:874-882

WILKS, D.S., (1995). Statistical Methods in the Atmospheric Sciences. International Geophysics Series, 59. Academic Press pp.465.

WILSON, L.J., (2000). Comments on “Probabilistic Predictions of Precipitation Using the ECMWF Ensemble Prediction System”. Weather and Forecasting. 15:361-364.